

# Characterizing Multistationarity Regimes in Biochemical Reaction Networks

Irene Otero-Muras<sup>1</sup>, Julio R. Banga<sup>2</sup>, Antonio A. Alonso<sup>2\*</sup>

**1** Department of Biosystems Science and Engineering, The Swiss Federal Institute of Technology Zurich, Zurich, Switzerland, **2** BioProcess Engineering Group, Instituto Investigaciones Marinas- Consejo Superior de Investigaciones Cientificas, Spanish National Research Council, Vigo, Spain

## Abstract

Switch like responses appear as common strategies in the regulation of cellular systems. Here we present a method to characterize bistable regimes in biochemical reaction networks that can be of use to both *direct* and *reverse* engineering of biological switches. In the design of a synthetic biological switch, it is important to study the capability for bistability of the underlying biochemical network structure. Chemical Reaction Network Theory (CRNT) may help at this level to decide whether a given network has the capacity for multiple positive equilibria, based on their structural properties. However, in order to build a working switch, we also need to ensure that the bistability property is robust, by studying the conditions leading to the existence of two different steady states. In the reverse engineering of biological switches, knowledge collected about the bistable regimes of the underlying potential model structures can contribute at the model identification stage to a drastic reduction of the feasible region in the parameter space of search. In this work, we make use and extend previous results of the CRNT, aiming not only to discriminate whether a biochemical reaction network can exhibit multiple steady states, but also to determine the regions within the whole space of parameters capable of producing multistationarity. To that purpose we present and justify a condition on the parameters of biochemical networks for the appearance of multistationarity, and propose an efficient and reliable computational method to check its satisfaction through the parameter space.

**Citation:** Otero-Muras I, Banga JR, Alonso AA (2012) Characterizing Multistationarity Regimes in Biochemical Reaction Networks. PLoS ONE 7(7): e39194. doi:10.1371/journal.pone.0039194

**Editor:** Johannes Jaeger, Centre for Genomic Regulation (CRG), Universitat Pompeu Fabra, Spain

**Received:** November 4, 2011; **Accepted:** May 20, 2012; **Published:** July 3, 2012

**Copyright:** © 2012 Otero-Muras et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Funding provided by Spanish government, Ministry of Science and Innovation project MultiSysBio (ref. DPI2008-06880-C03-02) and DPI2004-07444-C04-03, Xunta de Galicia (PGDIT02-PXIC40209PN). Consejo Superior de Investigaciones Cientificas intramural project BioREDES (ref. PIE-201170E018). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: antonio@iim.csic.es

## Introduction

Multistability is a nonlinear phenomenon characterized by the existence of two or more stable steady states, where a given dynamical system will evolve depending on its initial conditions. Important biological phenomena, like cellular decision processes, rely on multistable models, where the different functional phenotypic states or cell fates can be understood as *discrete, stable and mutually exclusive* stable states [1]. Experimental evidences for bistability have been found in numerous pathways involved in cell decision processes, such as the p42 MAPK/Cdc2 network governing the maturation of oocytes in *Xenopus* [2], the pheromone sensing MAPK pathway in *S. cerevisiae* [3], or the Rb-E2F pathway regulating proliferation in mammalian cells [4]. The analysis of mathematical models of the underlying multistable networks contributes to understand these biological phenomena from a systems perspective. In [5] for example, basic design principles for the control of the cell cycle are suggested based on the modeling and analysis of the gene circuit underlying the Rb-E2F switch, identified by the criterion of robustness.

The dynamics of biochemical reaction networks (i.e., the time evolution of the vector of species concentrations) can be described by models of coupled ordinary differential equations where the structure depends on the reaction connectivities, stoichiometry and kinetics, and the parameters are defined from kinetic rate

constants. Modeling a biochemical system consists of inferring the structure and parameters from experimental data. For a given model structure, the corresponding parameters are typically estimated from experimental time course measurements of observables (usually linear combinations of some subsets of the species concentrations), by minimizing some measure of error between the experimental data and the model prediction [6]. Nontraditional methodologies for the determination of reaction mechanisms from kinetic data sets have been reviewed by [7,8]. In processes occurring within living cells, the access to quantitative information is often very limited, and this fact has severe implications on the development of mathematical models: it hampers model discrimination, often leads to poor parameter identifiability, and makes the parameter estimation task very challenging, since it entails to solve a nonconvex optimization problem in high dimensional search spaces [9], which cannot be reduced and/or constrained in absence of a priori knowledge about feasible parameter values.

One of the challenges of systems biology is to provide tools to overcome the lack of quantitative information, by exploring and exploiting connections between model structure and/or parameters with the expected dynamic behaviour [10–12]. In this context, methods to systematically detect multistationary regimes in biochemical systems will help modeling multistable systems, constraining the feasible parameter regions, for a given structure,

based on the capability to produce multistationarity. These methods are also of great interest in the design of synthetic biological switches, where the robustness of the multistationarity property needs to be analyzed [13–15].

Current results in this direction are derived from different fields, from classical bifurcation theory [16] to monotone systems [17]. In particular, structural properties of reaction networks and their connection with multistationarity are at the core of the Chemical Reaction Network Theory, pioneered by Feinberg, Jackson and Horn [18,19] and subject to ongoing development [20,21] with special interest in the application to biological systems [22–24]. In the context of cell signaling, for example, CRNT has been used to discard kinetic mechanisms based on their capacity for multistationarity [25,26].

The deficiency one algorithm, the advanced deficiency theory, the deficiency zero and deficiency one theorems are part of the CRNT in which networks are classified by means of a nonnegative integer index called *deficiency*—a property of the graph of complexes of a network— and some structural conditions are evaluated to decide whether networks have the capacity for multiple positive equilibria [27].

As pointed out in [28], CRNT provides surprisingly strong results for reaction networks based only on the systems structure. For example, the deficiency zero theorem asserts that every weakly reversible network of zero deficiency has a unique equilibrium, for any choices of parameter values. However, when multistationarity cannot be ruled out, nothing can be said about how the parameters affect the qualitative behaviour of the solutions.

In a previous paper [29] we have introduced the parameters into the picture, providing a canonical expression for the equilibrium manifold in terms of the kinetic parameters and the so called deficiency parameters of the network. The concept of network *layout*, introduced in [29] as the difference between the deficiency of a network ( $\delta$ ) and the dimension of the equilibrium manifold ( $\lambda$ ), allowed us to classify biochemical reaction networks in three groups: *proper* networks ( $\delta = \lambda$ ), *overdimensioned* networks ( $\delta > \lambda$ ) and *underdimensioned* networks ( $\delta < \lambda$ ). The analysis presented focused on proper networks, i.e. those networks where deficiency and manifold dimension coincide. For those networks, the qualitative behaviour of the manifold was evaluated through its derivative with respect to the deficiency parameters. Then a geometric intuitive idea was applied to find, under these restrictive assumptions ( $\delta = \lambda$ ), a condition on the parameters of the network giving room to multiple steady states. The condition was formulated as a (non convex and multimodal) optimization problem, to be solved by global optimization algorithms. Before the search, it was needed to partition the parameter space in regions with different qualitative behaviour. The parameter space was then characterized depending on whether the optimization algorithm could find a solution or not, i.e., depending on whether or not we could find a point (or a set of discrete points) in the parameter space fulfilling the multistationarity condition.

In this work, we formalize the geometric intuition in [29] to derive a sufficient condition for multistationarity which is general, i.e. applies to networks where the dimension of the equilibrium manifold is lower, equal or greater than the deficiency. The general condition is stated in a formal context, and a proof of its validity in arbitrary dimensional spaces is provided. In addition, we present a method to search the condition through the parameter space with several advantages over global optimization methods. On the one hand, it is capable of finding the regions where multistationarity condition is fulfilled, without requiring the a priori partitioning of the parameter space in areas of different

qualitative behavior. On the other hand, the method allows the characterization of the multistationarity regimes of a biochemical reaction network in a reliable manner, i.e. those regions in the parameter-state space leading to multiple steady states are ensured to be enclosed by the solution set.

## Fundamentals

Following CRNT classical description (see Feinberg's Lecture 3 in [30] for details), we consider a generic reaction network involving  $m$  species  $\{\mathcal{S}_1, \dots, \mathcal{S}_m\}$  participating on a given set of irreversible reaction steps. Their concentrations  $c_i$  are collected on a vector  $\mathbf{c}$  defined on the space  $\mathbb{R}_{\geq 0}^m$  we will refer to as *the species space*. In the following, we write  $\mathbf{x} \in \mathbb{R}_{> 0}^m$  if  $x_i > 0$  for all  $i = 1, \dots, m$  and  $\mathbf{x} \in \mathbb{R}_{\geq 0}^m$  if  $x_i \geq 0$  for all  $i = 1, \dots, m$ . Each reaction step will be represented by an arrow which connects two particular combination of species, thus indicating how a given set of reactants is transformed into a certain set of products. The set of species at both extremes of the arrows are known in CRNT as *complexes*. The set of all complexes connected by reaction steps conforms the reaction network which is represented by a directed graph (C-graph or graph of complexes), where arrows (edges) indicate the reaction steps and the nodes correspond with the complexes (see Figure 2 for an example).

Let  $\{\mathcal{C}_1, \dots, \mathcal{C}_n\}$  be the set of complexes of the network. To each complex  $\mathcal{C}_i$  we associate a set  $I_i$  of integer elements which collects the indexes of those complexes that are directly reached from  $\mathcal{C}_i$  and a pair of vectors  $\{y_i, \mathbf{e}_i\}$ . The set  $I_i$  can be formally defined as:

$$I_i = \{j \in \{1, \dots, n\} | \mathcal{C}_j \text{ is directly reached from } \mathcal{C}_i\}.$$

Vector  $y_i \in \mathbb{R}_{\geq 0}^m$  contains the molecularities of the species in complex  $i$ , and  $\mathbf{e}_i$  is a vector of the standard basis of  $\mathbb{R}^n$  such that for every  $i, j = 1, \dots, n$ :

$$\varepsilon_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

The complete set of edges in the graph is constructed by connecting  $\mathcal{C}_i \rightarrow \mathcal{C}_j$  for all complexes  $i = 1, \dots, n$ . Every edge in the graph (or reaction step) directly linking complex  $\mathcal{C}_i$  to complex  $\mathcal{C}_j$ , has its corresponding reaction rate, of the form:

$$R_{ij} = k_{ij} \psi_i \quad (1)$$

where  $k_{ij} > 0$  is a constant parameter and  $\psi_i$  a scalar function:

$$\psi_i(\mathbf{c}) : \mathbb{R}_{\geq 0}^m \rightarrow \mathbb{R}_{\geq 0}. \quad (2)$$

In what follows we assume that the reaction rates are mass action. Thus each function  $\psi_i(\mathbf{c})$  takes the form:

$$\psi_i(\mathbf{c}) = \prod_{j=1}^m c_j^{y_i^j}. \quad (3)$$

Provided that  $\mathbf{c} \in \mathbb{R}_{> 0}^m$  (i.e. it is a strictly positive vector), the expression (3) can also be written as:

$$\ln \psi(\mathbf{c}) = Y^T \ln \mathbf{c}, \quad (4)$$

where the natural logarithm operator  $\ln(\cdot)$  acts on any vector element-wise.

The C-graph of a reaction network is composed by a number  $\ell$  of “isolated” sub-graphs known in CRNT as *linkage classes*  $\{L_1, \dots, L_\ell\}$ , (see Lecture 3, page 14 in [30] for a complete discussion), each containing a number of complexes  $\theta_k$  so that:

$$\sum_{k=1}^{\ell} \theta_k = n.$$

The reaction network presented in Figure 2 consists of two linkage classes: one involving complexes  $\{C_1, C_5\}$ , and the other involving complexes  $\{C_2, C_3, C_4\}$ . Each linkage class  $L_k$  is accompanied by a vector  $\mathbf{L}_k \in \mathbb{R}_{\geq 0}^n$  of the form:

$$\Lambda_{jk} = \begin{cases} 1 & \text{if } C_j \text{ in } L_k \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Complexes within a linkage class are linked by sequences of reaction steps (or equivalently by sequences of arrows) defining paths. Two complexes are strongly linked if they can be reached from each other by directed paths (every complex is strongly linked to itself). The maximal set of strongly linked complexes is a strong terminal linkage class if no other complex can be reached from any of its elements. A linkage class that is also a strong terminal linkage class is said to be *weakly reversible*. Weakly reversible networks are those composed of weakly reversible linkage classes. The network depicted in Figure 2 is an example of a weakly reversible network.

**The dynamics of chemical reaction networks.** Making use of the above definitions, the time evolution of species concentrations can be described by a set of ordinary differential equations of the form

$$\dot{\mathbf{c}} = \sum_{i=1}^n \sum_{j \in I_i} (\mathbf{y}_j - \mathbf{y}_i) R_{ij}. \quad (6)$$

This system can be rearranged into the more familiar form, extensively employed in the context of CRNT:

$$\dot{\mathbf{c}} = \mathbf{f}(\mathbf{c}) \equiv Y \cdot A[\psi(\mathbf{c})], \quad (7)$$

where vector  $\psi(\mathbf{c}) \in \mathbb{R}_{\geq 0}^n$  contains as entries the scalar functions  $\psi_i$  (2). In what follows, we will refer to the space  $\mathbb{R}_{\geq 0}^n$  in which every  $\psi$  is defined as *the space of complexes*. The mapping  $A: \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}^m$  is such that:

$$A(\psi) = \sum_{i=1}^n \psi_i(\mathbf{c}) \sum_{j \in I_i} k_{ij} \cdot (\varepsilon_j - \varepsilon_i), \quad (8)$$

and  $Y$  in (7) is the  $m \times n$  molecularity matrix, with columns being the molecularity vectors  $y_1, \dots, y_n$  of complexes  $C_1, \dots, C_n$ . The right hand side of equation (7)  $\mathbf{f}(\cdot): \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}^m$  is known in CRNT as the species formation function [18].

Trajectories of system (7) are constrained in the concentration space by some invariants of motion to lie on convex regions within the non-negative orthant known as reaction polyhedrons [31], or *stoichiometric compatibility classes* in the CRNT formalism (see Lecture 2, page 17, Definition 2.9 in [30]).

In order to characterize these regions, and suggested by the structure of the operator  $A$  (Eq. (8)), let us first consider the subspace spanned by vectors  $(\varepsilon_j - \varepsilon_i)$ , where the image of the operator  $A$  lies. The subspace is formally defined as:

$$\Delta = \text{span}\{(\varepsilon_j - \varepsilon_i) | i = 1, \dots, n \text{ and } j \in I_i\}. \quad (9)$$

It must be noted that the number of independent vectors  $(\varepsilon_j - \varepsilon_i)$  per linkage class is  $\theta_k - 1$ , and vectors from different linkage classes are independent, thus the dimension of  $\Delta$  is  $n - \ell$ . Related to  $\Delta$  there is another subspace we define as follows:

$$\Sigma = \text{span}\{(\mathbf{y}_j - \mathbf{y}_i) | i = 1, \dots, n \text{ and } j \in I_i\}. \quad (10)$$

The subspace  $\Sigma$  is known in CRNT as the *stoichiometric subspace*, where the species formation function  $\mathbf{f}(\mathbf{c})$  lies, since from (7) and (8) we also have that:

$$\mathbf{f}(\mathbf{c}) = \sum_{i=1}^n \psi_i(\mathbf{c}) \sum_{j \in I_i} k_{ij} \cdot (\mathbf{y}_j - \mathbf{y}_i). \quad (11)$$

Let  $s$  be the dimension of  $\Sigma$  and  $B \in \mathbb{R}^{(m \times m - s)}$  a full rank matrix that spans column-wise the orthogonal complement  $\Sigma^\perp$ .

A linear vector-valued function of the form  $W(\mathbf{c}) = B^T \mathbf{c}$  characterizes the invariants of motion, since it is constant along the trajectories defined by (7) for any initial condition  $\mathbf{c}(0) = \mathbf{c}_0$ . This can be shown by differentiating  $W(\mathbf{c})$  along (7) so that:

$$\dot{W} = B^T \mathbf{f}(\mathbf{c}) = 0,$$

where the last equality follows because by construction  $B$  is orthogonal to the stoichiometric subspace  $\Sigma$ . Integrating this expression in time and taking into account the initial condition  $\mathbf{c}_0$  we get that:

$$B^T (\mathbf{c} - \mathbf{c}_0) = 0.$$

Each equation in the above expression is called a *conservation law* [32]. In this way, trajectories are constrained to regions that result from the intersection of the non-negative orthant  $\mathbb{R}_{\geq 0}^n$  with any linear variety associated to the stoichiometric subspace. These regions we will refer to as reaction polyhedrons (equivalently stoichiometric compatibility classes), can be formally defined with respect to a reference concentration vector  $\mathbf{c}_0$  as:

$$\Omega(\mathbf{c}_0) = \{\mathbf{c} \geq 0 | B^T (\mathbf{c} - \mathbf{c}_0) = 0 \text{ where } \text{Im} B \equiv \Sigma^\perp\}. \quad (12)$$

Mass conservation is a special case of invariance which leads to a particular class of polyhedrons, namely those which intersect every axis in the concentration space (in this case every species

participate in at least one conservation law). Other polyhedrons are possible being parallel to some axis as it is the case of the example. In the extreme case (no species participating in conservation laws) the stoichiometric subspace spans the whole concentration space so that the orthogonal complement is the zero vector (the only column of  $B$ ). In this case the reaction polyhedron is  $\mathbb{R}_{\geq 0}^m$ .

**The nature of equilibrium points.** Next we summarize some results from CRNT to be used in the contribution concerning possible equilibrium solutions of (7), namely vectors  $\mathbf{c}^*$  such that  $\mathbf{f}(\mathbf{c}^*) = 0$ . In what follows, we will concentrate on *weakly reversible* networks which in addition, for any initial condition  $\mathbf{c} \geq 0$  produce equilibrium points in the interior of the positive orthant. Note that if the trajectories lie in the interior of the positive orthant the networks under study are *persistent* [32] (i.e. those which for any initial condition  $\mathbf{c} > 0$  produce trajectories lying in  $\mathbb{R}_{> 0}^m$ ).

Any equilibrium point  $\mathbf{c}^* \in \mathbb{R}_{> 0}^m$  for (7) will satisfy either:

$$\psi(\mathbf{c}^*) \in D_0 \equiv \ker A \quad (13)$$

or

$$A[\psi(\mathbf{c}^*)] \in \ker Y \text{ with } \psi(\mathbf{c}^*) \notin \ker A. \quad (14)$$

Related to condition (14) there exists a subspace  $D_\delta \equiv \ker Y \cap \text{Im} A$ , we will refer to as the *deficiency subspace*, which plays a central role in CRNT. Its dimension is called the deficiency  $\delta$  and can be computed by making use of the standard relationship between the dimensions of the domain, kernel and image of a linear transformation (see [30], Lecture 4, page 23 in the proof of Proposition 4.7). For weakly reversible networks, the terms in the relationship coincide with the dimensions of subspaces  $\Delta$  (see (9)),  $D_\delta$  and  $\Sigma \equiv \text{Im} YA$ , respectively. The final expression then reads:

$$\dim(\ker Y \cap \text{Im} A) + \dim(\text{Im} YA) = \dim(\Delta), \quad (15)$$

since from previous discussion  $\dim(\Delta) = n - \ell$  and  $\dim(\text{Im} YA) = s$ , we finally get for  $\dim(\ker Y \cap \text{Im} A)$ :

$$\delta = n - \ell - s. \quad (16)$$

It must be noted that as stated in the so-called *deficiency zero theorem* [30] (Theorem 5.1 in Lecture 5, page 2), any equilibrium point fulfilling (13) will be stable and unique in each compatibility class. Consequently, multistationarity, that is to say multiple equilibria, can only occur for  $\mathbf{c}^*$  satisfying (14). For this reason, any weakly reversible reaction network of zero deficiency possesses a unique and stable equilibrium point (associated to  $D_0$ ) per stoichiometric compatibility class. This result remains valid independently of the values taken by the reaction rate constants.

In a previous paper [29], we have exploited the graph structure of biochemical networks to obtain an expression of the locus of equilibria –the set of points  $\mathbf{c}^*$  such that  $\mathbf{f}(\mathbf{c}^*) = 0$  in (7)– in terms of as many parameters as the deficiency of the network. For a class of networks (the so called *proper networks*) we were able to partition (by continuation of variation parameters associated with the deficiency) the space of kinetic parameters in regions with different qualitative dynamic behavior. Exploiting the concept of equilibrium as a intersection between solutions satisfying (14) with the so called reaction polyhedron [31], conditions on the parameters of

the network leading to multiple steady states were found for networks where the manifold dimension and deficiency coincide.

Here we start from this insight to provide a general condition for the existence of multistationarity, valid for weakly reversible mass action networks.

## Analysis

### The Locus of Equilibria

We present in this section a canonical expression for the locus of equilibria, that is to say the set of all possible feasible equilibrium solutions in terms of the kinetic parameters of the network. In what follows, we will refer to this locus as the *equilibrium manifold*. Mathematically it corresponds with an algebraic variety which results from the intersection of two other varieties: the *family of solutions* and the *mass action manifold* to be described below.

In preparing for the description and for simplicity, we assume that the molecularity matrix  $\mathcal{Y}$  is full rank and  $m \leq n$ . Furthermore, we assume that the  $m$  independent molecularity vectors  $y_i$  are distributed among linkage classes so that each linkage has at least one independent vector. For such networks, and without loss of generality, let us number the first  $\ell$  complexes so that each belongs to a different linkage class (note that by previous assumption  $m \geq \ell$ ), and so that the first  $m$  columns of the matrix  $Y$  are linearly independent.

**The family of solutions** is a linear variety  $\mathcal{F}$  defined in the space of complexes by:

$$\mathcal{F} = \sum_{k=1}^{\ell} \alpha_k \psi_k + \sum_{j=1}^{\delta} \alpha_j \mathbf{f}_j \quad (17)$$

where  $\alpha_j$  are given real numbers, and vectors  $\mathbf{x}_k, \mathbf{f}_j \in \mathbb{R}_{\geq 0}^n$  are solutions of the following equations:

$$A(\mathbf{x}_k) = 0 \quad k = 1, \dots, \ell, \quad (18)$$

$$A(\mathbf{f}_j) = \mathbf{w}_j \quad j = 1, \dots, \delta. \quad (19)$$

The set of vectors  $\{\mathbf{w}_j\}_{j=1}^{\delta}$  that appears in (19), defines a basis for the deficiency subspace  $D_\delta$ . As proved in [29], the elements of a basis for  $D_\delta$  can be obtained from the left kernel of the matrix:

$$[A \quad Y^T] \quad (20)$$

where  $\mathcal{Y}$  is the molecularity matrix and  $A$  is the  $n \times \ell$  matrix with columns being the vectors  $\mathbf{A}_k$  defined in (5).

Vectors  $\mathbf{x}_k \in \mathbb{R}_{\geq 0}^n$  for  $k = 1, \dots, \ell$  constitute a basis for the kernel of  $A$  (Eq. (8)). As stated in Proposition 4.1 of [30] (Lecture 4, page 10), for weakly reversible networks the dimension of the kernel of  $A$ , and therefore the number of vectors of the basis, coincides with the number of linkage classes. Actually, the same holds for networks other than weakly reversible, provided that they have one terminal linkage class per linkage class. The same Proposition prescribes for each element of the basis a nonnegative vector. In particular, each vector  $\mathbf{x}_k$  associated to a linkage class  $L_k$ , will have entries of the form:

$$x_{ik} = \begin{cases} 1 & i = k, \\ \rho_{ik} > 0 & i \neq k, \quad C_i \text{ in } L_k \\ 0 & i \neq k, \quad C_i \text{ not in } L_k \end{cases} \quad (21)$$

where parameters  $\rho_{ik} > 0$  are functions of the reaction constants within the linkage class  $L_k$ . The functional relation will be formally represented as  $\rho_{ik}(\mathbf{K}) > 0$  where  $\mathbf{K}$  is the vector of kinetic rate constants. A proof of Proposition 4.1 is out of the scope of this contribution. It cannot be found in [30] either. However, two alternative proofs can be found in the literature: one based on Perron-Frobenius theorems applicable to weakly reversible networks [33]. The other more graph theoretically oriented has been proposed in [34].

Similarly, any entry  $f_{ij}(\mathbf{K})$  of vectors  $\mathbf{f}_j$  is a function of kinetic parameters within the linkage class containing the complex  $C_j$ . Explicit expressions relating parameters  $\rho_{ik}$  and  $f_{ij}$  (for  $i = 1, \dots, n$ ,  $j = 1, \dots, \delta$ ) with the original kinetic constants are obtained by solving Eqs. (18) and (19).

It must be noted that by construction,  $\mathcal{F}$  in Eq. (17) under linear transformation  $A$  produces vectors:

$$A\mathcal{F} = \sum_{j=1}^{\delta} \alpha_j \mathbf{w}_j \quad (22)$$

which correspond with elements of the deficiency subspace  $D_\delta$ . For the particular case of  $\alpha_i = 0$  for every  $i = 1, \dots, \delta$ ,  $\mathcal{F}$  characterizes elements of  $D_0$  (i.e. the kernel of  $A$ ). In this way, Eq. (17) provides a complete parametrization of complexes  $\psi$ , leading to equilibrium solutions.

**The mass action manifold** is a nonlinear algebraic variety  $\mathcal{M}$  defined in the space of complexes by:

$$\ln \mathcal{M} = Q \ln \psi_m \equiv \sum_{j=1}^m q_j \ln \psi_j \quad (23)$$

where matrix  $Q$  is of the form:

$$Q^T = Y_1^{-1} Y.$$

with  $Y_1$  being a matrix containing the first  $m$  columns of the molecularity matrix  $Y$ . Note that each element  $\mathcal{M}_i$  can be written as:

$$\mathcal{M}_i = \prod_{j=1}^m \psi_j^{q_{ij}}.$$

Expression (23) is the equivalent of (4), but defined in the space of complexes instead of the concentration space. The relationship between both spaces is given by the bijective mapping:

$$\ln \psi_m = Y_1^T \ln \mathbf{c}. \quad (24)$$

**The equilibrium manifold** is the algebraic variety which results from the intersection in the space of complexes of the family of solutions  $\mathcal{F}$  and the mass action manifold  $\mathcal{M}$ . Formally the intersection can be expressed as:

$$\mathcal{F}(\psi_m, \alpha; \mathbf{K}) - \mathcal{M}(\psi_m) = 0 \quad (25)$$

where  $\alpha$  is the vector that collects all the  $\alpha_i$  parameters for  $i = 1, \dots, \delta$ . For a given parameter vector  $\mathbf{K}$ , the equilibrium

manifold can be written as:

$$\mathcal{H}_c(\psi_m, \alpha; \mathbf{K}) = 0 \quad (26)$$

where

$$\mathcal{H}_c(\psi_m, \alpha; \mathbf{K}) : \mathbb{R}_{>0}^m \times \mathbb{R}^\delta \rightarrow \mathbb{R}^{n-\ell}. \quad (27)$$

On occasions it may be convenient to transform the algebraic variety by means of (24) to its equivalent in the concentration space, namely:

$$\mathcal{H}_s(\mathbf{c}, \alpha; \mathbf{K}) = 0, \quad (28)$$

where:

$$\mathcal{H}_s(\mathbf{c}, \alpha; \mathbf{K}) : \mathbb{R}_{>0}^m \times \mathbb{R}^\delta \rightarrow \mathbb{R}^{n-\ell}. \quad (29)$$

It should be noted that given a rate constant vector  $\mathbf{K}$ , function  $\mathcal{H}_s$  (or equivalently  $\mathcal{H}_c$ ) is continuous and differentiable. In addition the dimension of the expression (28) (or equivalently (26)), either in the space of complexes or in the species space is  $\lambda = m - s$ . This is so since  $\lambda = m + \delta - (n - \ell)$ , and by (16)  $\delta = n - \ell - s$ .

### Condition for Multistationarity

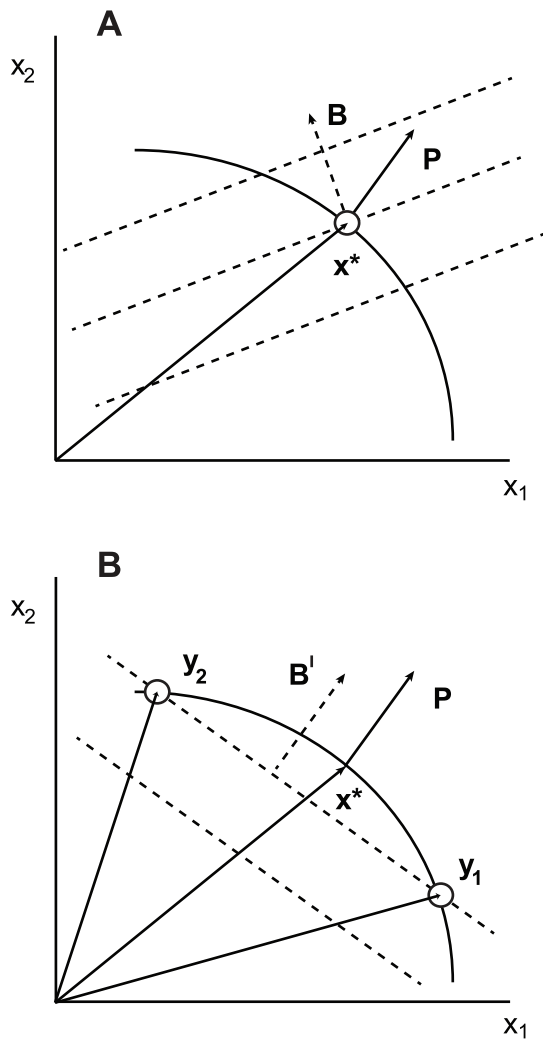
This section contains the main result of the contribution namely a condition for a given network to have multiple (positive) equilibria within the same stoichiometric compatibility class. The condition is geometric in essence and makes use of the equilibrium manifold, expressed either in the space of complexes (26) or in the species space (28), and the linear variety that defines the set of reaction polyhedrons (12).

The underlying idea behind the condition relates to the question of whether or not a given set of equations can accept more than one solution. A formal statement and discussion of this question is presented in the Appendix S1 for a general class of functions defined in  $\mathbb{R}^n$ . There we present the necessary mathematical background as well, but first let us illustrate the basic concept on a simple two dimensional case depicted in Figure 1.

The example consists of a nonlinear manifold (the continuous curve) and two possible families of linear varieties represented by dashed lines with gradient vectors  $\mathbf{B}$  and  $\mathbf{B}'$  respectively. Vector  $\mathbf{P}$ , perpendicular to the curve at  $x^*$ , is the gradient of the curve at  $x^*$ . This vector also defines the tangent subspace to the curve at that point. As it can be observed in Figure 1 A, all linear varieties associated to vector  $\mathbf{B}$  intersect the curve in just one point so no multiple solutions are expected.

On the other hand, some linear varieties associated to  $\mathbf{B}'$  in Figure 1 B intersect the curve in two points  $y_1$  and  $y_2$  what corresponds with two different solutions. What differs between Figures 1 A and 1 B is the relative orientations of the curve and linear variety gradients. Thereby vector alignment (or linear dependency) is what seems to be at stake to determine the number of solutions. In this way, multiple solutions are expected to appear whenever vector alignment takes place. This is the notion we take advantage of and extend to higher dimensional manifolds (i.e. the equilibrium manifold).

As discussed in the Appendix S1, essentially all is needed is the equilibrium manifold to be locally smooth. If at a given point in the space of concentrations this is the case, hyperplanes secant to



**Figure 1. Schematic representation of the geometric condition.**

Solid and dashed lines represent nonlinear and linear varieties, respectively in the two dimensional  $(x_1, x_2)$  space. Vector  $P$  is perpendicular to the curve (nonlinear variety) at  $x^*$ .  $B$  and  $B'$  in Figure (a) and (b) respectively, denote vectors perpendicular to the linear variety (straight lines defining compatibility classes). In Figure (a) vector independency leads to one intersection per compatibility class, while in Figure (b) vector alignment leads to two different intersections (indicated by  $y_1$  and  $y_2$ ).

doi:10.1371/journal.pone.0039194.g001

the equilibrium manifold can be constructed in a small neighborhood of the point by parallel translation of the corresponding hyperplanes tangent at that point (see Figure S1). Multiple solutions are then possible if the hyperplanes coincide with a given reaction polyhedron.

In the remaining of the section the results discussed in the Appendix S1 will be adapted to detect multiple equilibria associated to the equilibrium manifold. To that purpose a  $m + \delta$  (equivalently  $n - \ell + m - s$ ) dimensional space will be employed, which includes the variables:

$$(\mathbf{c}, \boldsymbol{\alpha}) = (c_1, c_m, \alpha_1, \alpha_\delta).$$

It must be pointed out that since the map (24) is bijective, the condition can be established either in the species space ( $c$  variables) or in

the space of complexes ( $\psi_m$  variables). However in the species space the reaction polyhedron is linear what simplifies the derivation of the condition. Thus for convenience, manifold (28) will be employed in first place. Nevertheless, some comments will be made at the end of the section on the condition expressed in the space of complexes.

Firstly let us note that the Jacobian of  $\mathcal{H}_s$  reads:

$$D\mathcal{H}_s = [D_c \mathcal{H}_s \ D_a \mathcal{H}_s] \quad (30)$$

where  $D_c \mathcal{H}_s \in \mathbb{R}^{(n-\ell) \times m}$  and  $D_a \mathcal{H}_s \in \mathbb{R}^{(n-\ell) \times \delta}$  denote the Jacobians of  $\mathcal{H}_s$  with respect to  $\mathbf{c}$  and  $\boldsymbol{\alpha}$ , respectively.

In the space described by the variables  $(\mathbf{c}, \boldsymbol{\alpha})$  we express the linear variety associated to the reaction polyhedron (12) as:

$$W(\mathbf{c}, \boldsymbol{\alpha}; \mathbf{c}_0) = 0 \text{ for every } \mathbf{c} \in \Omega(\mathbf{c}_0) \quad (31)$$

where function  $W : \mathbb{R}_{>0}^m \times \mathbb{R}^\delta \rightarrow \mathbb{R}^{m-s}$  is of the form:

$$W(\mathbf{c}; \mathbf{c}_0) = B^T (\mathbf{c} - \mathbf{c}_0). \quad (32)$$

Note that because  $D_c W = B^T$  and  $D_a W = 0$  (i.e. a zero  $(m-s) \times \delta$  matrix), its Jacobian can be written as:

$$DW = [B^T \ 0]. \quad (33)$$

$DW$  is full rank by construction, since as discussed in the Fundamentals,  $B$  is a basis for the orthogonal complement of the stoichiometric subspace  $\Sigma$ .

For a given vector of rate constants  $\mathbf{K}$ , let  $\mathcal{H}_s(\mathbf{c}, \boldsymbol{\alpha}; \mathbf{K}) = 0$  be continuous in the vicinity of a point  $(\mathbf{c}^*, \boldsymbol{\alpha}^*)$ . Note that by the implicit function theorem, this implies that  $D\mathcal{H}_s(\mathbf{c}^*, \boldsymbol{\alpha}^*; \mathbf{K})$  is full rank. Then we are under the conditions of Proposition A1 (see Appendix S1) where  $\mathcal{H}_s$  corresponds with  $F(x)$ ,  $x \equiv (\mathbf{c}, \boldsymbol{\alpha})$  and  $P = D\mathcal{H}_s^T(\mathbf{c}^*, \boldsymbol{\alpha}^*; \mathbf{K})$ . Furthermore,  $DW^T$  takes the place of  $C$  in matrix  $G$  in Corollaries A1 and A2. The corresponding  $n - \ell + m - s$  square  $G$  matrix then becomes:

$$G(\mathbf{c}^*, \boldsymbol{\alpha}^*; \mathbf{K}) = \begin{bmatrix} D_c \mathcal{H}_s & D_a \mathcal{H}_s \\ B^T & 0 \end{bmatrix}. \quad (34)$$

We are now in the position to formally state the geometric condition.

**Proposition 1** Consider a reaction network with a given vector of rate constants  $\mathbf{K}$ , and let  $\mathcal{H}_s(\mathbf{c}, \boldsymbol{\alpha}; \mathbf{K}) = 0$  be continuous in its domain. If for any  $(\mathbf{c}, \boldsymbol{\alpha})$  satisfying  $\mathcal{H}_s(\mathbf{c}, \boldsymbol{\alpha}; \mathbf{K}) = 0$  matrix  $G$  is full rank, the reaction network for  $\mathbf{K}$  has at most one positive equilibrium solution per stoichiometric compatibility class.

**Proof:** the result follows directly from Corollary A1 (see Appendix S1), applied to the domain where the equilibrium manifold (29) is defined (i.e. positive concentration space).

**Proposition 2** Given a vector of rate constants  $\mathbf{K}$ , a sufficient condition for the reaction network to exhibit multiple (positive) steady states within the same stoichiometric compatibility class is that for at least some  $(\mathbf{c}^*, \boldsymbol{\alpha}^*)$  such that  $\mathcal{H}_s(\mathbf{c}^*, \boldsymbol{\alpha}^*; \mathbf{K}) = 0$ , matrix  $G(\mathbf{c}^*, \boldsymbol{\alpha}^*; \mathbf{K})$  is rank deficient. Equivalently, a sufficient condition for the reaction network to exhibit multistationarity is that:

$$\det[G(\mathbf{c}^*, \boldsymbol{\alpha}^*; \mathbf{K})] = 0 \quad (35)$$

for some  $(\mathbf{c}^*, \boldsymbol{\alpha}^*)$ .

**Proof:** The result follows directly from Corollary A2 (see Appendix S1), applied to the domain where the equilibrium manifold (29) is defined (i.e. positive concentration space). Rank deficiency can be checked through expression (35).

The condition for multistationarity remains valid in the space of complexes, since the map (24) is bijective. In this space however,  $W(\psi_m; \mathbf{c}_0) = 0$  is nonlinear, although continuous for every  $\psi_m \in \mathbb{R}_{>0}^m$  (namely in the interior of the space of complexes). This can be shown by using (24) to compute its Jacobian with respect to  $\psi_m$ , so that:

$$D\psi_m W = \sigma^T B \quad (36)$$

with  $\sigma^T$  being of the form:

$$\sigma = \text{diag}(\mathbf{c}) \cdot (Y_1^T)^{-1} \cdot \text{diag}^{-1}(\psi_m) \quad (37)$$

where  $\text{diag}(\mathbf{v})$  and  $\text{diag}^{-1}(\mathbf{v})$  are the diagonal and inverse diagonal matrices operating over the vector  $\mathbf{v}$ , respectively.

The Jacobian  $D\psi_m W$  is full rank since  $B$  is full rank and  $\sigma \in \mathbb{R}^{m \times m}$  invertible. Continuity of  $W(\psi_m; \mathbf{c}_0) = 0$  then follows from the implicit function theorem (see Appendix S1). In the space of complexes, Condition (35) from Proposition 2 should be checked on the matrix:

$$G(\psi_m^*, \alpha^*; \mathbf{K}) = \begin{bmatrix} D\psi_m \mathcal{H}_c & D\alpha \mathcal{H}_c \\ B^T \sigma & 0 \end{bmatrix}. \quad (38)$$

## Interval Based Search

In order to find the regions in the parameter space fulfilling the multistationarity condition we formulate a so called *continuous constraint satisfaction problem* (CSP) [35,36] and solve it numerically by using interval methods [37]. Methods based on interval analysis allow mathematical operations to be carried out over real intervals instead of real numbers, and thus to represent a continuum of solutions to a given CSP by a finite number of intervals or boxes, the union of which encloses the solution set.

Let  $\mathbf{z} \in \mathbb{R}^p$  be the space where a set of constraints is defined. A domain in that space is constructed by interval variables  $z_i$  (for  $i = 1, \dots, p$ ) defined on closed real intervals  $[z_{il}, z_{iu}]$ , where subindexes  $l$  and  $u$  stand for lower and upper bounds, respectively. For  $p$  variables, the cartesian product of  $\diamond$ -interval domains  $\mathbf{z} = z_1 \times \dots \times z_p$  is called a box.

Following [36], a constraint satisfaction problem consists in finding an interval domain where a set of constraints  $\mathcal{G} = g_1, \dots, g_q$  hold. This can be formally stated as:

$$\begin{cases} g_j(z_1, \dots, z_p) \diamond 0 & j = 1, \dots, q \\ z_i = [z_{il}, z_{iu}] & i = 1, \dots, p \end{cases}$$

where constraints involve nonlinear analytic expressions and the symbol  $\diamond$  stands for either equality or inequality constraints, that is to say  $\diamond \in \{=, \leq, \geq\}$ . A solution of a CSP is an element of the search space which fulfills all the equalities and inequalities simultaneously.

In our case, the search space involves the elements of the rate constant vector  $\mathbf{K}$  and the independent variables which characterize the equilibrium manifold (which equals  $\lambda = m - s$ , the manifold dimension). In this way, for  $\mathbf{K} \in \mathbb{R}_{>0}^{\rho}$ , the number of variables is  $p = \rho + m - s$ .

Regarding constraints, the equality ones correspond with the equilibrium manifold plus condition (35) which adds up to  $n - \ell + 1$  equations. On the other hand, inequality constraints must be imposed on the dependent variables that describe the equilibrium manifold, to ensure positivity of variables representing concentration and to search on non-zero  $\alpha$  values.

Characterizing multistationarity regimes boils down to find the regions in the parameter-state space composed of the rate constants and independent variables which fulfill the constraints, i.e. computing all real feasible solutions of the corresponding CSP.

Solutions in interval methods are approximated by *subpavings* which consist of unions of boxes. Formally, a subpaving  $\mathbb{S}$  for a given region  $R$  of the search space is defined as the union of nonoverlapping boxes approximating  $R$ . If we construct subpavings  $\underline{\mathbb{S}}$  and  $\overline{\mathbb{S}}$  such that:

$$\underline{\mathbb{S}} \subset R \subset \overline{\mathbb{S}},$$

the region  $R$  is bracketed between inner and outer approximations. The outer approximation is reliable [38], since it is guaranteed that the solution region is contained within  $\overline{\mathbb{S}}$ .

Simple algorithms as SIVIA (Set Inverter Via Interval Analysis) proposed in [37] can be used to compute inner and outer subpavings by successive bisections and selections. Implementing interval algorithms requires environments supporting interval arithmetics such as the free available software package INTLAB, which provides an interactive environment within Matlab [39,40]. Efficiency can be gained by branch-and-prune algorithms where a set of boxes that contains all the solutions of the CSP is computed, and then each box is reduced and split. The free software package REALPAVER [36] provides a modeling language and a generic branch and prune algorithm which combine different splitting strategies and pruning techniques.

## Example: Multistationary Regimes in an Autocatalytic Network

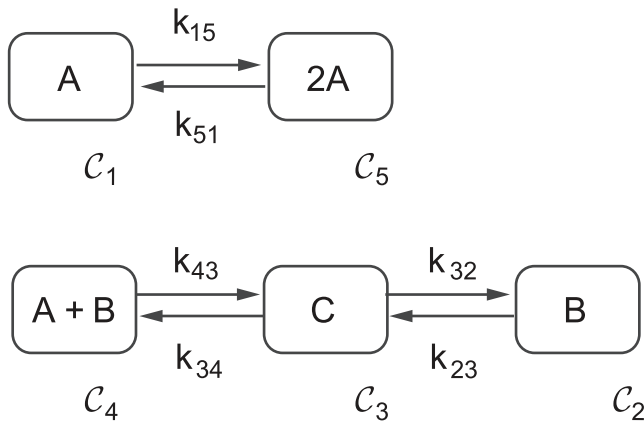
As a proof of concept, we make use of a simple autocatalytic network introduced by Edelstein [41], and already used as a case study for the bifurcation analysis in the context of biochemical systems [42]. Bistability has been shown for very simple autocatalytic systems like phosphorylation-dephosphorylation cycles with autocatalytic kinase [43].

The Edelstein network involves three species  $\{A, B, C\}$ , distributed into five complexes,  $\{A, B, C, A+B, 2A\}$ . The C-graph for the Edelstein network is depicted in Figure 2, where the following index sets indicate complex interconnections:  $\mathcal{I}_1 = \{5\}$ ,  $\mathcal{I}_2 = \{3\}$ ,  $\mathcal{I}_3 = \{2, 4\}$ ,  $\mathcal{I}_4 = \{3\}$ ,  $\mathcal{I}_5 = \{1\}$ . The reaction steps that correspond with the set of edges in the graph are presented in Table 1. The network has two different linkage classes:  $L_1$  containing the complexes  $(\mathcal{C}_1, \mathcal{C}_5)$ , and  $L_2$  containing the complexes  $(\mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4)$ . Their corresponding vectors  $\mathbf{\Lambda}_1$  and  $\mathbf{\Lambda}_2$  (see (5)) read:

$$\mathbf{\Lambda}_1 = (1 \ 0 \ 0 \ 0 \ 1)$$

$$\mathbf{\Lambda}_2 = (0 \ 1 \ 1 \ 1 \ 0).$$

From the C-graph structure it follows that the network is weakly reversible since each linkage class is weakly reversible (in fact the network is reversible).



**Figure 2. Graph of complexes (C-Graph) for the Edelstein network.** The network involves three chemical species  $A$ ,  $B$  and  $C$ . Nodes in the graph correspond with complexes, denoted by  $C_j$  with  $j=1, \dots, 5$ . Arrows in the graph represent reaction steps. In this example every reaction is reversible. doi:10.1371/journal.pone.0039194.g002

The species space is in this case three dimensional ( $m=3$ ) with concentrations  $c_1=[A]$ ,  $c_2=[B]$  and  $c_3=[C]$ . The molecularity matrix  $Y$  reads:

$$Y = \begin{bmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

with rank equal to the number of species, i.e.  $\text{rank}(Y)=3$ . The dynamics of the reaction network can be encoded by Eq. (7), where the expression for  $A[\psi]$  reads:

$$A[\psi] = \begin{bmatrix} -k_{15} & 0 & 0 & 0 & k_{51} \\ 0 & -k_{23} & k_{32} & 0 & 0 \\ 0 & k_{23} & -(k_{32}+k_{34}) & k_{43} & 0 \\ 0 & 0 & k_{34} & -k_{43} & 0 \\ k_{15} & 0 & 0 & 0 & -k_{51} \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \\ \psi_5 \end{bmatrix}.$$

The elements of vector  $\psi$  (mass action monomials) are related to concentrations by Eq. (3) so that:

$$\psi = (c_1 \quad c_2 \quad c_3 \quad c_1 c_2 \quad c_1^2)^T.$$

The time evolution of species concentrations results from substituting previous expressions for  $Y$ ,  $A[\psi]$  and  $\psi$  into Eq. (7) so that:

$$\dot{c}_1 = k_{15}c_1 - k_{51}c_1^2 + k_{34}c_3 - k_{43}c_1c_2$$

$$\dot{c}_2 = -k_{23}c_2 + (k_{32}+k_{34})c_3 - k_{43}c_1c_2$$

$$\dot{c}_3 = k_{23}c_2 - (k_{32}+k_{34})c_3 + k_{43}c_1c_2.$$

**Table 1. Edelstein network reaction steps.**

Reaction step	Rate
$A \rightarrow 2A$	$R_{15} = k_{15} \cdot [A]$
$2A \rightarrow A$	$R_{51} = k_{51} \cdot [A]^2$
$A + B \rightarrow C$	$R_{43} = k_{43} \cdot [A][B]$
$C \rightarrow A + B$	$R_{34} = k_{34} \cdot [C]$
$C \rightarrow B$	$R_{32} = k_{32} \cdot [C]$
$B \rightarrow C$	$R_{23} = k_{23} \cdot [B]$

doi:10.1371/journal.pone.0039194.t001

The stoichiometric subspace  $\Sigma$  defined in (10) is in this case:

$$\Sigma = \text{span}\{\mathbf{y}_1 - \mathbf{y}_5, \quad \mathbf{y}_2 - \mathbf{y}_3, \quad \mathbf{y}_3 - \mathbf{y}_4\}$$

being its dimension  $s=2$  (i.e. two linearly independent vectors). The corresponding matrix  $B$  (with columns defining a basis of the orthogonal complement  $\Sigma^\perp$ ) is:

$$B^T = [0, 1, 1].$$

For reference concentrations  $c_{10}$ ,  $c_{20}$ ,  $c_{30}$  the linear variety (31) is a plane of the form:

$$c_2 + c_3 - c_{20} - c_{30} = 0.$$

According to formula (16) we have that  $\delta = 5 - 2 - 2 = 1$ . Thus the deficiency subspace  $D_\delta$  is one dimensional with the following basis computed from the left kernel of matrix (20):

$$\mathbf{w} = \begin{bmatrix} 1 & -1 & 0 & 1 & -1 \end{bmatrix}^T.$$

For this network, the expression for the family of solutions (17) takes the form:

$$\mathcal{F} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \rho_{51} \end{bmatrix} \psi_1 + \begin{bmatrix} 0 \\ 1 \\ \rho_{32} \\ \rho_{42} \\ 0 \end{bmatrix} \psi_2 + \alpha \begin{bmatrix} 0 \\ 0 \\ -f_3 \\ -f_4 \\ +f_5 \end{bmatrix},$$

where the explicit expressions relating the parameters  $\rho_{ij}$  and  $f_{i1}$  (subindex 1 is omitted), obtained by solving Eqs. (18) and (19), are given in Table 2.

In order to compute the mass action manifold (23) we first select matrix  $Y_1$  as that which collects the first  $m=3$  molecularity vectors:



**Table 2.** Edelstein network parameter equivalences.

Parameter	Equivalence
$\rho_{51}$	$\frac{k_{15}}{k_{51}}$
$\rho_{32}$	$\frac{k_{23}}{k_{32}}$
$\rho_{42}$	$\frac{k_{23}k_{34}}{k_{32}k_{43}}$
$f_3$	$\frac{1}{k_{32}}$
$f_4$	$\frac{k_{32}+k_{34}}{k_{32}k_{43}}$
$f_5$	$\frac{1}{k_{51}}$

doi:10.1371/journal.pone.0039194.t002

$$Y_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Matrix  $Y_1$  defines the bijective mapping (24). The resulting mass action manifold becomes:

$$\ln \mathcal{M} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 2 \end{bmatrix} \ln \psi_1 + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \ln \psi_2 + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \ln \psi_3.$$

Using (25), we compute the equilibrium manifold as the intersection of the family of solutions with the mass action manifold. Its dimension is  $\lambda = m - s = 1$  and its expression in the space of complexes (26) is given by the equations:

$$\rho_{32}\psi_2 - \alpha f_3 - \psi_3 = 0$$

$$\rho_{42}\psi_2 - \alpha f_4 - \psi_1\psi_2 = 0$$

$$\rho_{51}\psi_1 + \alpha f_5 - \psi_1^2 = 0.$$

The bijective mapping (24) allows us to write the equations for the equilibrium manifold in the species space (28):

$$\rho_{32}c_2 - \alpha f_3 - c_3 = 0 \quad (39)$$

$$\rho_{42}c_2 - \alpha f_4 - c_1c_2 = 0 \quad (40)$$

$$\rho_{51}c_1 + \alpha f_5 - c_1^2 = 0. \quad (41)$$

In order to derive the geometric condition defined by (35) in Proposition 2, we first compute the Jacobian of  $\mathcal{H}_s$  as described in (30), where:

$$D_c \mathcal{H}_s = \begin{bmatrix} 0 & \rho_{32} & -1 \\ -c_2 & \rho_{42} - c_1 & 0 \\ \rho_{51} - 2c_1 & 0 & 0 \end{bmatrix} \text{ and } D_\alpha \mathcal{H}_s = \begin{bmatrix} -f_3 \\ -f_4 \\ +f_5 \end{bmatrix}.$$

Matrix  $G$  in (34) becomes:

$$G = \begin{bmatrix} 0 & \rho_{32} & -1 & -f_3 \\ -c_2 & \rho_{42} - c_1 & 0 & -f_4 \\ \rho_{51} - 2c_1 & 0 & 0 & f_5 \\ 0 & 1 & 1 & 0 \end{bmatrix},$$

being its determinant:

$$\det(G) = -f_5c_2(1 + \rho_{32}) -$$

$$(2c_1 - \rho_{51})[f_4(1 + \rho_{32}) + f_3(c_1 - \rho_{42})].$$

Setting  $\det(G) = 0$  as in (35) and using the equivalences in Table 2, we get:

$$0 = (k_{23} + k_{32})k_{43}c_2 +$$

$$(2k_{51}c_1 - k_{15})(k_{32} + k_{34} + k_{23} + k_{43}c_1). \quad (42)$$

In the interior of the species space (i.e. that of positive concentrations), the first term at the right hand side of (42) is always positive, while the sign of the second term is conditioned by that of  $(2k_{51}c_1 - k_{15})$ . Therefore, the determinant can only vanish for values of  $c_1$  that make  $(2k_{51}c_1 - k_{15}) < 0$ . Equivalently, for  $c_1$  values within the open interval  $(0, \rho_{51}/2)$ , i.e.  $c_1$  satisfying:

$$c_1 = \frac{\beta \rho_{51}}{2} \text{ for some } 0 < \beta < 1.$$

Substituting this expression for  $c_1$  into (41) and its result in (40) we obtain, respectively:

$$\alpha = \frac{k_{15}^2 \beta (\beta - 2)}{4k_{51}},$$

and

$$c_2 = \frac{\beta(\beta-2)\rho_{51}^2 f_4}{2f_5(2\rho_{42} - \beta\rho_{51})}.$$

Since  $\beta \in (0,1)$  the term  $\beta-2 < 0$ . Thus for  $c_2$  to be positive it is required that:

$$2\frac{\rho_{42}}{\rho_{51}} < \beta < 1.$$

By substituting the expressions obtained for  $\alpha$  and  $c_2$  in the equilibrium manifold  $\mathcal{H}_s$ , it can be deduced that  $c_3$  is also positive for every  $\beta$  in the open interval  $(2\rho_{42}/\rho_{51}, 1)$ . Summarizing, the network will show multiplicities for those parameters fulfilling (42) with:

$$c_1 = \frac{\beta\rho_{51}}{2}, c_2 = \frac{\beta(\beta-2)\rho_{51}^2 f_4}{2f_5(2\rho_{42} - \beta\rho_{51})}$$

for  $\beta \in (2\rho_{42}/\rho_{51}, 1)$ . Note that multistationarity is ruled out for  $\frac{\rho_{42}}{\rho_{51}} > \frac{1}{2}$  since  $\beta < 1$ .

Let us consider the set of kinetic constants given in Table 3. For  $\beta = 0.6451$  they satisfy the condition described by (42). The resulting values for  $c_1$  and  $c_2$  are:

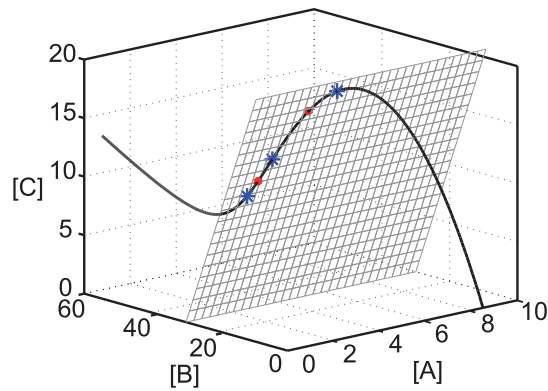
$$c_1 = 2.7416, c_2 = 12.4230.$$

In Figure 3, the equilibrium manifold for these values of the kinetic constants is depicted, together with the reaction polyhedron corresponding to  $[B]_0 + [C]_0 = 30$ . The manifold is one dimensional, and intersects the reaction polyhedron in three points, corresponding to three different equilibria. The points fulfilling the rank deficiency condition, corresponding to  $\alpha = -15.787$  and  $\alpha = -9.079$  are also indicated. As it can be deduced from the figure, three steady states will exist for a range of the sum of initial concentrations  $[B]_0 + [C]_0$ . In fact, performing a continuation of the curve of equilibria by varying the values of  $[B]_0 + [C]_0$  we obtain the curve shown in Figure 4, where two limit points or saddle node bifurcations appear for  $[B]_0 + [C]_0 = 29.7768$  and  $[B]_0 + [C]_0 = 30.6949$ . Within these values, corresponding to different positions of the reaction polyhedron, three steady states will exist. Note that the points fulfilling the rank deficiency

**Table 3. Edelstein Network parameters.**

Kinetic constant	Value
$k_{15}$	8.5
$k_{51}$	1
$k_{23}$	0.2
$k_{32}$	1
$k_{34}$	1
$k_{43}$	1

doi:10.1371/journal.pone.0039194.t003



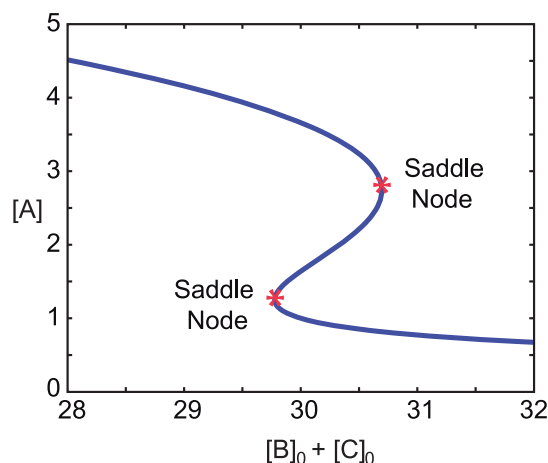
**Figure 3. Equilibrium manifold for the Edelstein network.** The corresponding parameter values are given in Table 3. Stars indicate steady states of the system for  $[B]_0 + [C]_0 = 30$ . Dots indicate those points in the locus of equilibria where the rank deficiency condition is fulfilled.

doi:10.1371/journal.pone.0039194.g003

condition indicated in Figure 3 correspond precisely with the bifurcation points in Figure 4.

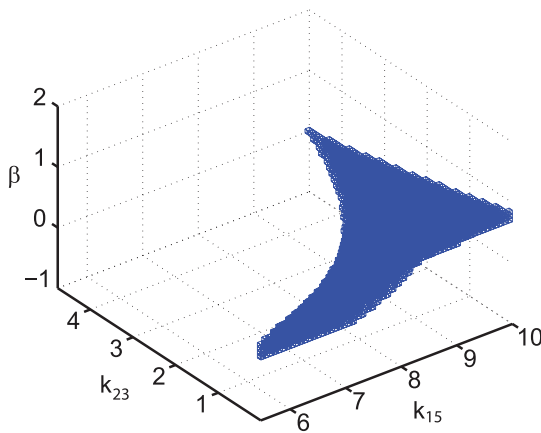
Alternatively, interval methods can be employed to search for the condition given by (35). Here it is important to remark that the method allows searching for parameters and/or steady state values of species concentrations within multistationary regimes, provided some other parameter values and/or steady state concentrations fixed. To illustrate this, let us assume we are interested in the ranges of parameters allowing for multistationarity, and the steady state values of  $c_1$ , for a given steady state concentration of the species  $B$  (i.e.  $[B] = c_2$ ). The variables in the constraint satisfaction problem are thus the free kinetic parameters and the steady state concentration of the species  $A$ .

A three dimensional plot of the result is given in Figure 5 for  $c_2 = 12.4230$ , where the  $x$  and  $y$  axis represent two of the kinetic parameters ( $k_{23}$  and  $k_{51}$ ), and the  $z$  axis represents a function of the steady state concentration of the species  $A$  in the steady state. We represent  $\beta$  in the  $z$  axis in order to facilitate the comparison with the results derived analytically. The corresponding values of  $c_1$  can be computed as  $c_1 = \beta\rho_{51}/2$ .



**Figure 4. Equilibrium curve for the Edelstein network.** The curve is obtained by varying  $[B]_0 + [C]_0$ , using the software CI Matcont [46]. The kinetic parameters are kept fixed with the values shown in Table 3.

doi:10.1371/journal.pone.0039194.g004



**Figure 5. Outer approximation of the region of multistationarity for the Edelstein Network.** The concentration  $[B]$  is fixed ( $c_2 = 12.4230$ ),  $k_{23} \in (0, 10)$ ,  $k_{15} \in (0, 10)$ , and values of  $k_{51}$ ,  $k_{32}$ ,  $k_{34}$  and  $k_{43}$  given in Table 3.  
doi:10.1371/journal.pone.0039194.g005

## Discussion

In this work we present a method to compute the regimes of multiple steady states in biochemical reaction networks, i.e. the regions in the parameter space, or in the state-parameter space of network models leading to multistationarity.

The main result of the paper consists of a sufficient condition for multistationarity, demonstrated to be valid for weakly reversible networks of arbitrary dimension and deficiency.

The idea behind is based on the fact that one steady state is an intersection between the locus of equilibria, or equilibrium manifold, and the reaction simplex, or stoichiometric compatibility class. In search for the existence of multiple steady states in reaction networks, we explored the requirements for multiple intersections between these two varieties.

In a previous work of the authors [29] it was shown that, for networks with  $\delta = \lambda$ , (or equivalently with  $m = n - l$ ), and fixed a kinetic rate vector  $\mathbf{K}$ , the equilibrium manifold could be continued by the variation of the deficiency parameters of the network, and its qualitative behaviour evaluated through the derivative of the manifold with respect to the deficiency parameters. After partitioning of the parameter space in regions of different qualitative behaviour, global optimization algorithms [44] were used to check, within every region, whether multiple intersections between the manifold and the simplex were possible.

Here we use and extend this geometric insight to state a general condition of multistationarity for networks where the deficiency might be lower, equal or greater than the dimension of the equilibrium manifold. In this regard, [29] deals with a particular case of the general condition presented (and proved) here.

The evaluation of the multistationarity condition presented boils down to check the rank of a matrix which depends on the kinetic parameters, the concentrations, and the deficiency parameters. This matrix is systematically derived as indicated in the Analysis section, from the equations of the manifold and the mass conservation laws. In order to check the condition through the (state-)parameter space, we reformulate the problem as a constraint satisfaction one to be solved by interval methods.

For the purpose of characterizing the (state-)parameter space in terms of the capability for multiple steady states, methods based on interval arithmetics presents several advantages over classical global optimization methods [44]. On the one hand, they allow

identifying regions, and not sets of discrete points, in the (state-)parameter space. In this way, there is no need to partition a priori the (state-)parameter space in regions with different qualitative behaviour. On the other hand, they ensure reliability of the solution [38], i.e. they guarantee that all the multistationarity regimes are enclosed by the solution regions.

As it has been commented in the Introduction, the CRNT provides particularly strong results to rule out multiple steady states based on the network structure irrespective of the network parameters. In this regard, this paper extends the results of the CRNT by introducing the kinetic parameters into the picture, giving a general condition for the appearance of multiple steady states.

In terms of applicability and performance, the method presented here is valid independently of the value of the network deficiency and dimension. The efficiency of the search will depend on the computational cost of the algorithm which increases in high dimensional spaces. In case we are only interested in finding points in the state-parameter space leading to multistationarity (for example if the goal is to decide only whether the network can exhibit multiple steady states or not), global optimization methods [44,45] would perform much faster.

Once the condition for multistationarity is derived, and depending on the particular scenario we want to explore, the search can be performed considering some of the parameters and/or states to be fixed. In this way we obtain bifurcation diagrams in the desired low dimensional projections of the state-parameter spaces much more efficiently than using standard continuation techniques. For example, if we are interested in computing all the parameter sets giving multiple steady states for a particular value of the equilibrium concentration of one the species, we can perform the search keeping this concentration fixed.

To end up, we would like to stress the applicability of our method to standard problems appearing in systems and synthetic biology. The method can be particularly convenient, for example, to evaluate the robustness of the switching property for a given network. As remarked by [13], building a synthetic switch does not only require to start from a network structure allowing bistability, but such bistability must be sufficiently robust, meaning that the range of parameters leading to bistability is wide enough to enable a successful practical implementation. In this context, our method will efficiently characterize the regions of interest within the state-parameter space.

## Supporting Information

**Figure S1 Representation of the tangent and secant hyperplanes associated to a nonlinear manifold, in this case a curve, at  $x^*$ ;  $e_1$  and  $e_2$  are unit vectors centered at  $x^*$ ;  $y_1$  and  $y_2$  are two different points of the manifold.**  
(EPS)

**Appendix S1**  
(PDF)

## Acknowledgments

We thank the anonymous reviewers for their helpful comments and suggestions.

## Author Contributions

Conceived and designed the experiments: IOM AAA. Performed the experiments: IOM. Analyzed the data: IOM AAA. Contributed reagents/materials/analysis tools: IOM AAA JRB. Wrote the paper: IOM AAA JRB.

## References

- Huang S (2005) Multistability and Multicellularity: Cell fates as high dimensional attractors of gene regulatory networks. In: Kriete A, Eils R, editors. *Computational Systems Biology*. Elsevier Academic Press.
- Xiong W, Ferrell JE (2003) A positive-feedback-based bistable memory module that governs a cell fate decision. *Nature* 426: 460–465.
- Paliwal S, Iglesias PA, Campbell K, Hilioti Z, Groisman A, et al. (2007) MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast. *Nature* 446: 46–51.
- Yao G, Lee TJ, Mori S, Nevins JR, You L (2008) A bistable Rb-E2F switch underlies the restriction point. *Nat Cell Biol* 10: 476–482.
- Yao G, Tan C, West M, Nevins JR, You L (2011) Origin of bistability underlying mammalian cell cycle entry. *Mol Syst Biol* 7: 485.
- Chen WW, Niepel M, Sorger PK (2010) Classical and Contemporary approaches to modeling biochemical reactions. *Genes Dev* 24: 1861–1875.
- Ross J, Vlad MO (1999) Nonlinear kinetics and new approaches to complex reaction mechanisms. *Annu Rev Phys Chem* 50: 51–78.
- Ross J, Schreiber I, Vlad MO (2006) Determination of complex reaction mechanisms: analysis of chemical, biological, and genetic networks. Oxford: Oxford University Press.
- Banga JR (2008) Optimization in computational systems biology. *BMC Syst Biol* 2: 47.
- Kitano H (2002) Systems biology: A brief overview. *Science* 295: 1662–1664.
- Schreiber I, Ross J (2003) Mechanism of oscillatory reaction deduced from bifurcation diagrams. *J Phys Chem A* 107: 9846–9859.
- Wang Y, Zhang XS, Chen L (2011) Computational systems biology: integration of sequence, structure, network, and dynamics. *BMC Syst Biol* 5: S1.
- Cherry JL, Adler FR (2000) How to make a biological switch. *J Theor Biol* 203: 117–133.
- Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403: 339–342.
- Tyson JJ, Albert R, Goldbeter A, Ruoff P, Sible J (2008) Biological switches and clocks. *J R Soc Interface* 5: S1–8.
- Lu J, Engl HW, Schuster P (2006) Inverse bifurcation analysis: application to simple gene systems. *Algorithms Mol Biol* 1: 1–11.
- Angeli D, Ferrell JE, Sontag ED (2004) Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems. *Proc Natl Acad Sci U S A* 101: 1822–1827.
- Horn FJM, Jackson R (1972) General Mass Action Kinetics. *Arch Rational Mech Anal* 47: 81–116.
- Feinberg M and Horn FJM (1974) Dynamics of open chemical systems and the algebraic structure of the underlying network. *Chem Eng Sci* 29: 775–787.
- Craciun G, Feinberg M (2005) Multiple equilibria in complex chemical reaction networks: I. The injectivity property. *SIAM J Appl Math* 65: 1526–1546.
- Craciun G, Feinberg M (2006) Multiple equilibria in complex chemical reaction networks: II. The species reaction graph. *SIAM J Appl Math* 66: 1321–1338.
- Craciun G, Tang Y, Feinberg M (2006) Understanding bistability in complex enzyme-driven reaction networks. *Proc Natl Acad Sci U S A* 103: 8697–8702.
- Conradi C, Flockerzi D, Raisch J, Stelling J (2007), Subnetwork analysis reveals dynamic features of complex (bio)chemical networks. *Proc Natl Acad Sci U S A* 104: 19175–19180.
- Shinar G, Feinberg M (2010), Structural sources of robustness in biochemical reaction networks. *Science*, 327: 1389–1391.
- Conradi C, Saez-Rodriguez J, Gilles E, Raisch J (2005) Using Chemical Reaction Network Theory to discard a kinetic mechanism hypothesis. *IEE Proc Syst Biol* 152: 243–248.
- Saez-Rodriguez J, Hammerle-Fickinger A, Dalal O, Klamt S, Gilles ED, Conradi C (2008) On the multistability of signal transduction motifs. *IET Syst Biol* 2(2):80–93.
- Feinberg M (1987) Chemical Reaction Network Structure and the Stability of Complex Isothermal Reactors I. The deficiency zero and deficiency one theorems. *Chem Eng Sci* 42 (10):2229–2268.
- Bailey JE (2001) Complex biology with no parameters. *Nat Biotechnol* 19: 503–504.
- Otero-Muras I, Banga JR, Alonso AA (2009) Exploring multiplicity conditions in enzymatic reaction networks. *Biotechnol Prog* 25: 619–631.
- Feinberg M (1979), Lectures on chemical reaction networks. Notes of lectures given at the Mathematics Research Center, University of Wisconsin. Available: [www.che.eng.ohio-state.edu/~einberg/LecturesOnReactionNetworks](http://www.che.eng.ohio-state.edu/~einberg/LecturesOnReactionNetworks).
- Otero-Muras I, Szederkenyi G, Alonso AA, Hangos KM (2008) Local dissipative Hamiltonian description of reversible reaction networks. *Syst Control Lett* 57: 554–560.
- Angeli D, de Leeuwe P, Sontag ED (2007) A Petri net approach to the study of persistence in chemical reaction networks. *Math Biosci* 210: 598–618.
- Horn FJM (1972) Necessary and sufficient conditions for complex balancing in chemical kinetics. *Arch Rational Mech Anal* 49: 172–186.
- Feinberg M, Horn FJM (1977) Chemical mechanism structure and the coincidence of the stoichiometric and kinetic subspaces. *Arch Rational Mech Anal* 66: 83–97.
- Pedamallu CS, Kumar A, Csendes T, Posfaj J (2011) An interval partitioning algorithm for constraint satisfaction problems. *Int J Model Ident Control* 14: 133–140.
- Granvilliers L, Benhamou F (2006) Reapaver: An interval solver using constraint satisfaction techniques. *ACM Trans Math Softw* 32: 138–156.
- Jaulin L, Kieffer M, Didrit O, Walter E (2001) Applied interval analysis. London: Springer.
- Moore RE (1988) Reliability in computing: the role of interval methods in scientific computing. San Diego: Academic Press Professional, Inc.
- Rump S M (1999) INTLAB - INTerval LABoratory, In: Tibor Csendes Editor. *Developments in Reliable Computing*. Dordrecht: Kluwer Academic Publishers. 77–104. Software Available: <http://www.ti3.tu-harburg.de/rump>.
- Pal L and Csendes T (2009) INTLAB implementation of an interval global optimization algorithm. *Optim Methods Softw* 24: 749–759.
- Edelstein B (1970) Biochemical model with multiple steady states and hysteresis. *J Theor Biol*, 29: 57–62.
- Chickarmane V, Paladugu SR, Bergmann F, Sauro HM (2005) Bifurcation discovery tool. *Bioinformatics* 21: 3688–3690.
- Bishop LM, Qian H (2010) Stochastic bistability and bifurcation in a mesoscopic signaling system with autocatalytic kinase. *Biophys J* 98: 1–11.
- Csendes T, Pal L, Sendin JOH, Banga JR (2008) The GLOBAL Optimization Method Revisited. *Optim Lett* 2: 445–454.
- Egea JA, Rodriguez-Fernandez M, Banga JR and Marti R (2007) Scatter search for chemical and bioprocess optimization. *J Global Optim* 37: 481–503.
- Dhooge A, Govaerts W, Kuznetsov YA, Mestrom W, Riet AM (2003) CL MATCONT: A continuation toolbox in Matlab. *Proceedings of the 2003 ACM Symposium on Applied Computing* 3: 161–166.